

Large-Scale Analysis of Drought in Europe Using Ncep/Ncar and Era-40 Re-Analysis Data Sets

I. Bordi¹, K. Fraedrich², M. Petitta¹ and A. Sutera¹

¹Department of Physics, University of Rome "La Sapienza", P.le Aldo Moro 2 I-00185 Rome, Italy

²Meteorologisches Institut, Universität Hamburg, D-20146 Hamburg, Germany

Abstract: Drought variability in Europe has been analysed using the two re-analysis data sets currently available: the National Center for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR) and the European Centre for Medium-Range Weather Forecasts (ECMWF, 40-yr re-analysis ERA-40). The aim was to compare the performances of these data sets in assessing the space-time variability of dry/wet periods at large scale. Drought conditions in the last forty years have been evaluated by computing the Standardized Precipitation Index (SPI), while the climatic variability has been analyzed through the Principal Component Analysis applied to the index time series. Results for the European sector show that the two re-analyses differ remarkably both in the first loading and in representing the timing of wet and dry periods. The first principal component scores are characterized by a linear trend superposed to short-term fluctuations. The analysis extended to the globe shows noticeable differences in the first loadings, while the two re-analyses agree in their first principal component scores, which are dominated by a long-term linear trend. Thus, both data sets capture a linear trend, but the areas where this feature should be observed are not uniquely identified.

Key words: Drought variability, SPI, long-term trend, re-analysis data.

1. INTRODUCTION

The currently available re-analysis products offer the opportunity to investigate the long-term and global scale variability of climate because they consist of self-consistent, comprehensive global data sets for long periods. For this purpose, the re-analysed data are extensively used world-wide. The analysis proposed here is based on the National Centers for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR) data set and the European Centre for Medium-Range Weather Forecasts (ECMWF) 40-yr re-analysis (ERA-40).

The NCEP/NCAR reanalyses, which are available back to 1948, contain several meteorological variables in a global spatial resolution and in a vertical extend from the surface toward the 10 hPa level. They are a composite of different data sources such as observations from land stations and ships, upper air rawinsondes, satellite and numerical weather forecasts, which are assimilated in an AGCM (Atmospheric Global Circulation Model) and re-analysed by means of a "frozen" state of an AGCM (for more information see Kalnay et al., 1996). The ECMWF 40-year re-analysis project (ERA-40, Simmons and Gibson, 2000, Uppala, 2002) has been recently finalised and the data are available for climatic studies. This analysis of the state of the atmosphere, which covers the period September 1957 to August 2002, complements the hitherto available NCEP/NCAR and ERA-15 re-analyses. The ERA-40 project applies a modern Variational Data Assimilation technique (used in daily operational numerical forecasting at ECMWF) to the past conventional and satellite observations.

In the present paper we study the large-scale drought variability in the last forty years using the two re-analysis data, putting particular attention on the possible existence of a linear trend in the climatic signal. This is to evaluate the impact of different spatial resolutions and data assimilation schemes, characterizing the re-analysis products, on the assessment of drought. At this aim we decide to use the Standardized Precipitation Index (SPI) as an indicator of dry/wet periods. This

index, in fact, because of its standardization is a useful tool for comparing the climatic conditions of areas characterized by different hydrological regimes. The study is along the line of several investigations on drought assessment in different areas of the world based on the application of the SPI (Bordi and Sutera, 2001, Bonaccorso et al., 2003, Bordi et al., 2004a, 2004b).

The paper is organised as follows. In section 2 data and methods for the analysis are presented, while in section 3 the main results for the Euro-Mediterranean area, obtained by applying the Principal Component Analysis (PCA) to the SPI time series, are shown. In section 4 the analysis for the two data sets is extended to the globe and in the final section some conclusions and discussions are put forward.

2. DATA AND METHODS OF ANALYSIS

The analysis is focussed on time series of precipitation fields, from which drought is monitored.

First, the two data sets analysed have different spatial resolutions and different assimilation schemes have been used in the re-analysis procedures. The ERA-40 precipitation data are interpolated to a $2.5^\circ \times 2.5^\circ$ regular latitude/longitude grid, while those from the NCEP/NCAR data have $1.9^\circ \times 1.9^\circ$ resolution. The ECMWF re-analysis is provided every six hours and there is the possibility to get the total precipitation or its different components (say convective and stratiform), while daily or monthly precipitation data are available for the NCEP/NCAR re-analysis. Because the two data sets cover different periods, we choose the common one ranging from January 1958 to August 2002. As a first step, we carry out the analysis for the Euro-Mediterranean region; the area selected is 12.5°W to 62.5°E , 27.5°N to 70°N . Next, we extend the study to the globe.

Secondly, drought conditions in the area of interest have been assessed by computing the SPI. The index was developed by McKee et al. (1993) to monitor drought on multiple time scales and compare climatic conditions of different areas having different hydrological regimes. Furthermore, it permits to assess wet periods at the same time and needs for its computation only long time series of monthly precipitation. Thus, it appears to be a useful tool for our purpose. The index is computed by fitting a probability density function to the frequency distribution of precipitation summed over the selected time scale (1, 2, 3 months, etc.). This is performed separately for each month of the year and each location in space. Each probability density function is then transformed into the standardised normal distribution, which readily allows comparison between distinct locations and analytical computation of exceedance probabilities. Values of the standardised normal variable are grouped into classes that identify the severity level of a drought or wet event (Bordi and Sutera, 2001). The index may be computed for different time scales allowing addressing the various types of droughts: the shorter time scales for meteorological and agricultural drought, the longer ones for hydrological drought, etc. The time scale of 24 months is a suitable time scale for the investigation of long-term aspects of dryness and wetness (see Keyantash and Dracup, 2002), thus we focus our analysis on this time scale.

As an example of the index application to the two data sets, we show in Figure 1 the time series of the SPI on 24-month time scale (hereafter SPI-24) for Izmir (Turkey) and Cyprus. Figures show encouraging similarities when the SPI-24 time series for Izmir is considered, while several differences are evident for Cyprus.

Thirdly, time and spatial variability of drought has been also investigated applying the Principal Component Analysis to the SPI-24 data. The PCA is a widely used data reduction statistical method generating a set of linearly independent spatial patterns (loadings). They display spatial variability in terms of statistically deduced spatial patterns, which describe the correlation between the SPI-series at single stations (grid point in our case) and the corresponding scores (principal component time series, PCs). These spatial patterns are ordered with respect to their contribution to the total variance denoted in percent (details are given by Rencher, 1998, see also Bordi and Sutera, 2001). To reduce the degrees of freedom we may select only few spatial patterns to represent the statistical properties of the original field. However, there is no general rule to decide how many eigenvectors

should be retained, though few methods have been proposed (Rencher, 1998). For our case, we decided to keep 10 loadings since they capture about 76% of the total variance.

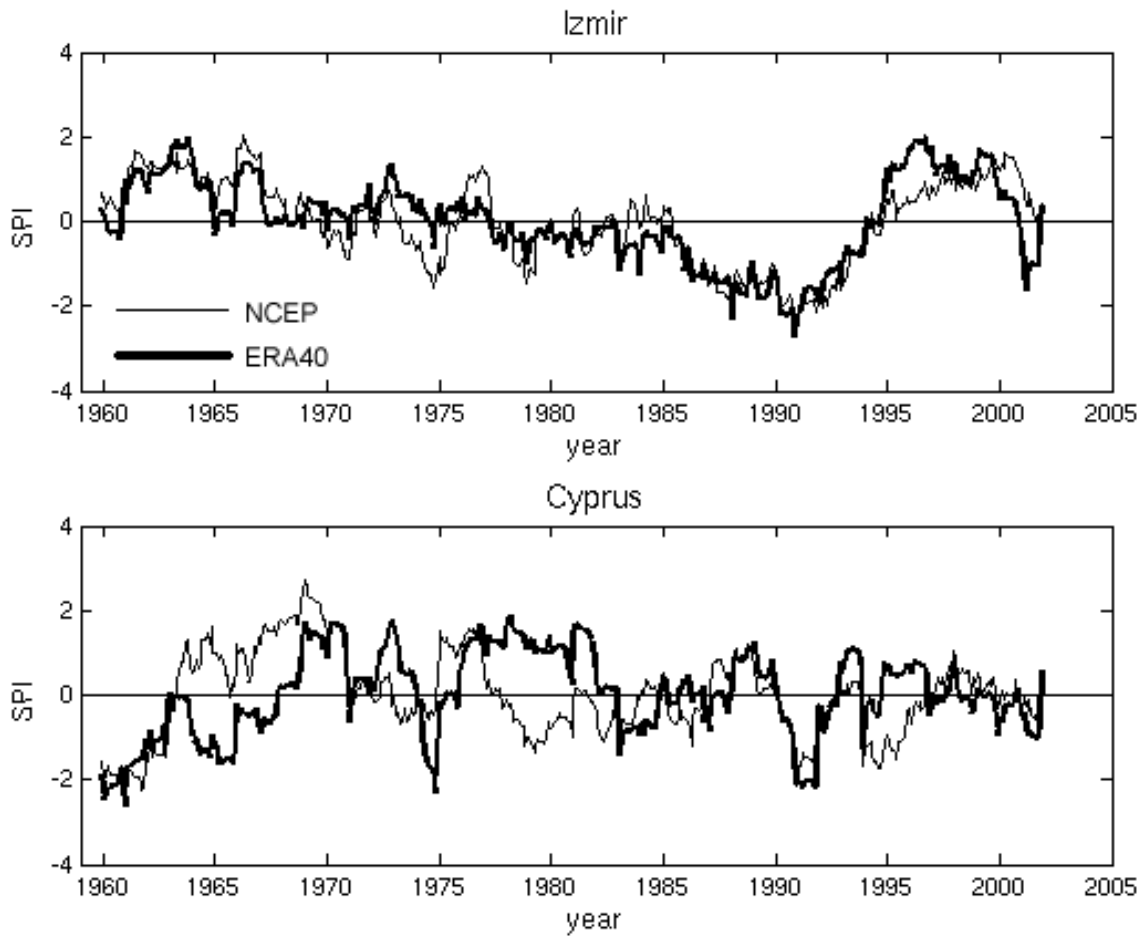


Figure 1. SPI-24 time series for Izmir (Turkey) and Cyprus. Data used are ERA-40 (thick line) and NCEP/NCAR (thin line).

3. PCA OF THE SPI-24: EURO-MEDITERRANEAN AREA

The PCA of the SPI-24 is applied to the Euro-Mediterranean area. The percentages of variance, for the two re-analysis data sets, explained by each loading are listed in Table 1. The first two loadings obtained using the ERA-40 precipitation data and the corresponding principal component scores are shown in Figure 2. The leading spatial pattern, which explains 22.9% of the total variance, shows positive correlation between the SPI time series and the corresponding score in the Balkans, Italy, central Europe and Spain, and negative correlations elsewhere. The first PC score (PC1) denotes a quick downhill towards negative values till the middle seventies and then for about ten years the PC time series remains negative. Up to middle 1980s the PC score is around the zero value, this means that all the regions that show high/low correlation with the first score time series in the last fifteen years experienced near normal conditions (see yellow/blue areas in loading 1).

The second loading, which explains 15.7% of the total variance, has instead a typical north-south structure (the pattern changes sign at about 45° to 50°N within the domain). It resembles the NAO (North Atlantic Oscillation) pattern though the associated score (PC2) differs from the time behaviour of the NAO index. The remaining loadings, which explain (each one) a little percentage of the total variance (see Table 1), don't have relevant spatial features and the corresponding PC scores show multi-year fluctuations.

Table 1. The first ten variance contributions of PC scores of the SPI-24 obtained using ERA-40 and NCEP/NCAR re-analysis data (second and third column respectively). The cumulative variance in both cases is about 76%.

n.	% variance ERA-40	% variance NCEP/NCAR
1	22.90	21.66
2	15.72	13.17
3	9.48	8.40
4	6.60	6.84
5	5.02	6.02
6	4.19	5.64
7	3.56	4.80
8	3.38	3.77
9	2.62	3.43
10	2.47	2.94
% Cumulative variance		76.67

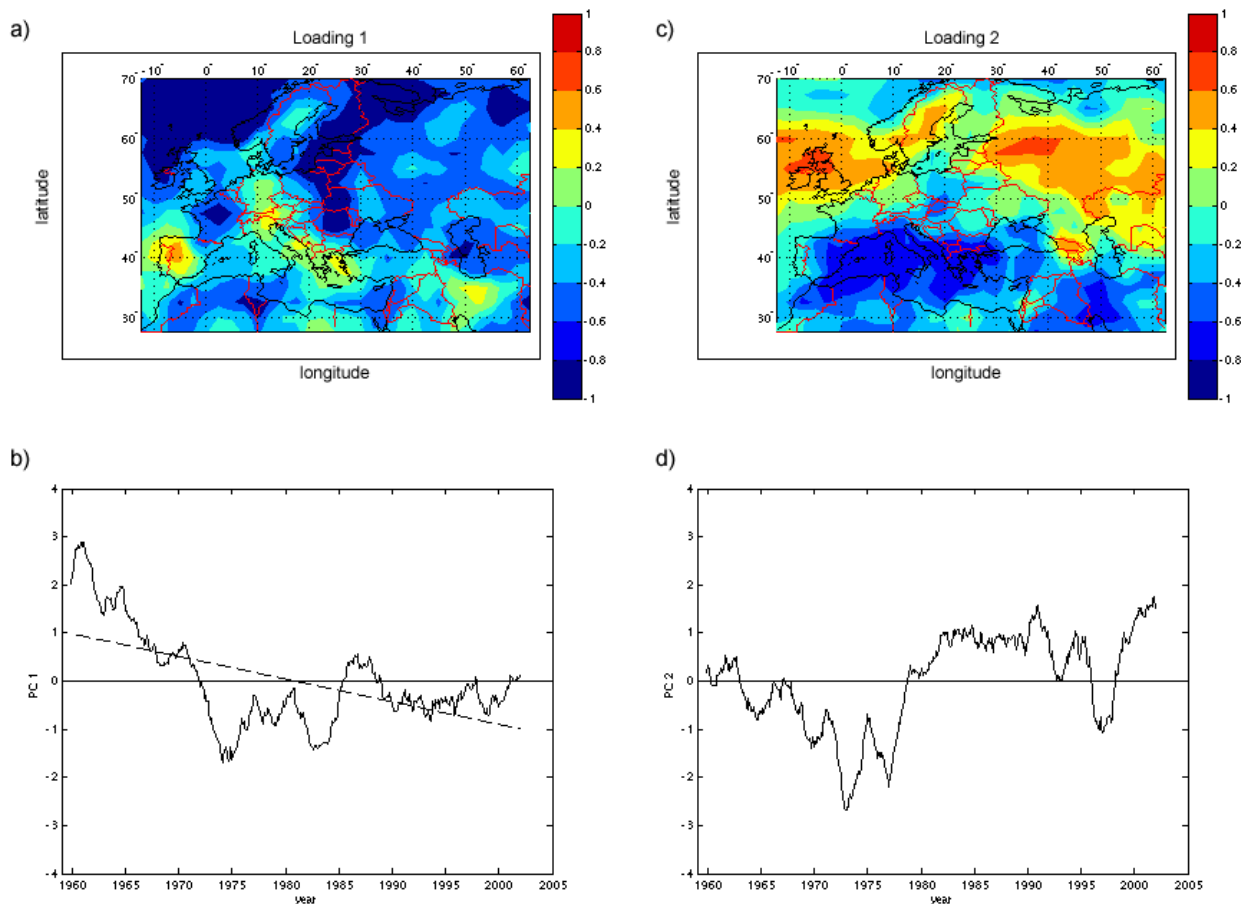


Figure 2: First two loading patterns, (a) and (c), and the associated principal component scores, (b) and (d), of the SPI-24 for the Euro-Mediterranean area. Data are monthly precipitation from ERA-40 data set. Contour interval is 0.2.

Results obtained with the NCEP/NCAR precipitation data sets are shown in Figure 3 (i.e. the first two loadings and corresponding PC scores). The first spatial pattern, which explains 21.7% of the total variance, reaches maximum values in North Africa, central Spain, north-eastern Europe, part of Italy, Balkans, Greece and Middle Orient. The time behaviour of the associated PC score reveals a trend towards negative values, meaning that the regions mentioned above from the middle seventies onward have been more affected by dry events. The second loading mode explains 13.2% of the total variance and has, as in the previous case, a north-south structure. Also the associated score is close to the one obtained with the ERA-40 data set.

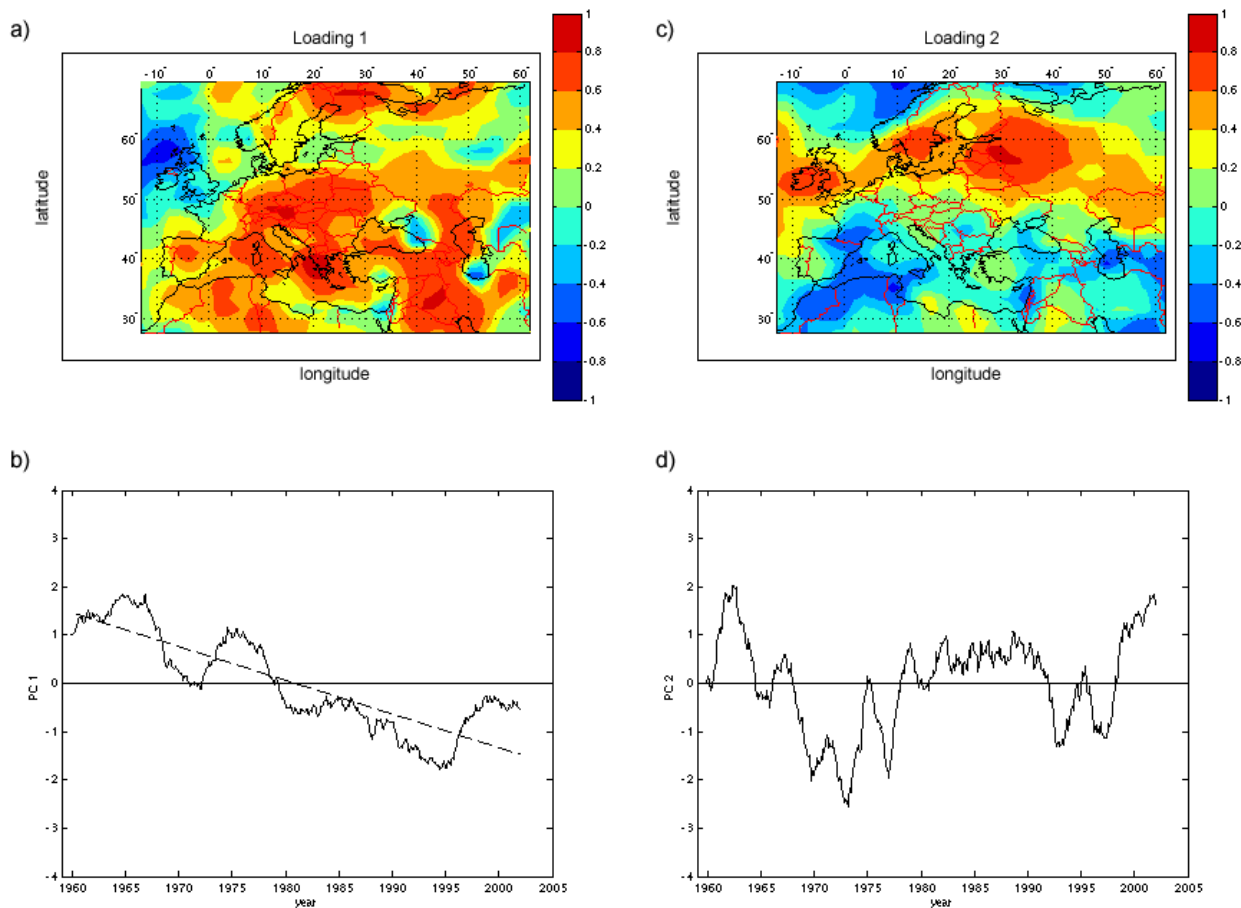


Figure 3: First two loading patterns, (a) and (c), and the associated principal component scores, (b) and (d), of the SPI-24 for the Euro-Mediterranean area. Data are monthly precipitation from NCEP/NCAR data set. Contour interval is 0.2.

Thus, we may argue that the two re-analysis data sets give remarkable differences in the first spatial pattern that describes the maximum variance of the SPI-24 field. Some discrepancies are also detectable between the two first PC scores. However, both scores are characterised by a long-term linear trend (see dashed lines in Figure 2b and 3b) superimposed to short-term fluctuations. Such fluctuations reduce the degree of co-variability between the two signals (correlation coefficient of 0.53).

In order to understand whether these differences are due to the coarse spatial resolution of the ERA-40 precipitation field, we employed the kriging procedure to this data set and repeated the analysis (say, computing the SPI-24 and PCA). Various kinds of interpolation have been applied (linear, quadratic, cubic). The analysis provides similar results compared to those based on the original precipitation data, suggesting that the origin of the detected differences cannot be attributed to the spatial resolution.

4. GLOBAL SCALE

We have extended the analysis to the globe in order to investigate whether the differences between the two precipitation data sets affect only the Euro-Mediterranean region or discrepancies are also detectable in other regions. The first loadings obtained for the two data sets and the corresponding PC scores are shown in Figure 4. There are several differences between the two first loadings, in particular in Europe, Asia, Northern America and in the Inter-Tropical Convergence Zone (ITCZ). The first PC scores for the two data sets, instead, show similar time behaviour

(correlation coefficient of 0.96) and denote a long-term linear trend towards negative values commencing in the seventies. That is, the orange/red areas in the loadings patterns have been affected by dry conditions in the last thirty years.

It should be noted that, due to the differences unveiled in the spatial patterns, the two data sets do not lead to the same drought assessment in some regions, such as Canada or Eastern Asia. We feel that to overcome this dichotomy and better understand the origin of the unveiled discrepancies a strictly comparison with the observations should be done, at least for sample areas.

5. CONCLUSIONS

Impacts on the assessment of drought variability of different spatial resolutions and data assimilation systems, characterising the currently available re-analysis data sets NCEP/NCAR and ERA-40, have been studied. We assessed the large-scale drought conditions over the Euro-Mediterranean area by computing the Standardized Precipitation Index over 24 months time scale. Then, applying the Principal Component Analysis to the SPI time series, the space-time climate variability in the last forty years has been analysed.

Results suggest that there are many differences between the two analyses that cannot be attributed to the different spatial resolutions of the two precipitation fields. In particular, for the European sector the two re-analyses differ remarkably both in the first loading and in representing the timing of wet and dry periods. The first principal component scores are characterized by a long-term linear trend superposed to short-term fluctuations.

When extending the analysis to the globe, the first spatial patterns of variability show different features. The corresponding PC scores show instead a good agreement and both are characterised by a long-term linear trend. This means that the areas where such trend should be observed are not uniquely identified by the two re-analyses.

The unveiled discrepancies are probably due to the different assimilation schemes adopted to produce the re-analysis products; thus, a stringent comparison with the rain gauge observations would be useful to understand the nature of the problem.

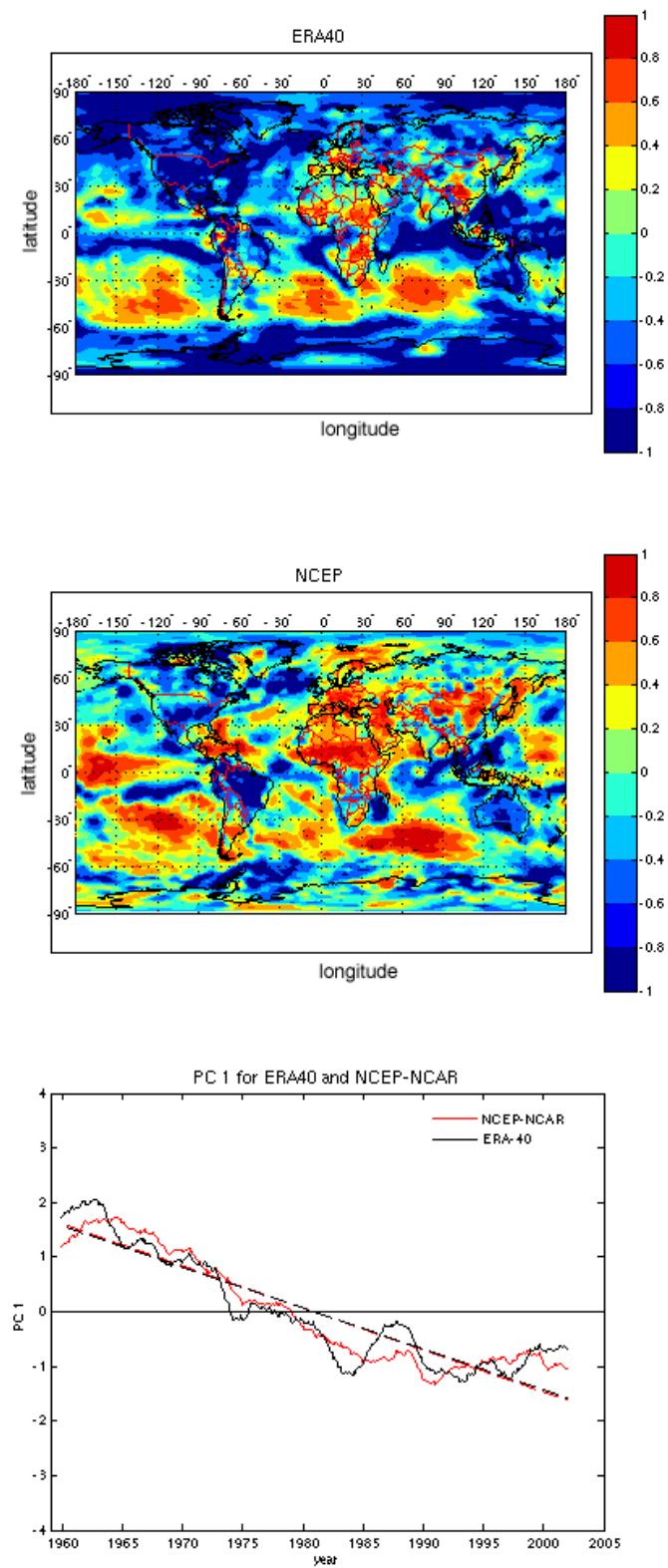


Figure 4: First loading pattern of the SPI-24 for the globe: a) ERA-40 b) NCEP/NCAR. Contour interval is 0.2 and the explained percentages of variances are 28.2% and 18.2% respectively. In figure c) there are the corresponding PC scores (black line refers to ERA-40 and red line to NCEP/NCAR).

ACKNOWLEDGMENT

NCEP/NCAR and CPC US Unified Precipitation data have been provided by the NOAA-CIRES Climate Diagnostics Center, Boulder, Colorado, USA, from their Web site at <http://www.cdc.noaa.gov>. ERA-40 precipitation data set has been provided by the European Centre for Medium-Range Weather Forecasts from their web site <http://www.ecmwf.int>.

REFERENCES

- Bonaccorso B., I. Bordi, A. Cancelliere, G. Rossi and A. Sutera, 2003: Spatial variability of drought: an analysis of the SPI in Sicily, *Water Res. Man.*, 17, 273-296.
- Bordi I., A. Sutera, 2001: Fifty years of precipitation: some spatially remote teleconnections, *Water Res. Man.*, 15, 247-280.
- Bordi I., K. Fraederich, F.-W Gerstengarbe, P. C. Werner and A. Sutera, 2004a: Potential predictability of dry and wet periods: Sicily and Elbe-Basin (Germany), *Theor. Appl. Climatol.*, 77, 125-138.
- Bordi I., K. Fraederich, J.-M. Jiang, A. Sutera, 2004b: Spatio-temporal variability of dry and wet periods in eastern China, *Theor. Appl. Climatol.*, 79, 81-91.
- Kalnay, E. and co-authors, 1996: The NCEP-NCAR 40 year reanalysis project. *Bull. Amer. Meteor. Soc.*, 77, 437-471.
- Keyantash, J. and J.A. Dracup, 2002: The quantification of drought: an evaluation of drought indices. *Bull. Amer. Meteor. Soc.*, 83, 1167-1180.
- McKee, T. B., N. J. Doesken, and J. Kleist, 1993: The relationship of drought frequency and duration to time scales. Preprints, 8th Conference on Applied Climatology, 17-22 January, Anaheim, CA, pp. 179-184.
- Rencher A. C., 1998: Multivariate statistical inference and applications. John Wiley & Sons, INC., 559 pp.
- Simmons, A.J., and J.K. Gibson, 2000: The ERA-40 project plan. ERA-40 Project Report Series 1, ECMWF, Reading, U.K., 62 pp.
- Uppala, S., 2002: ECMWF Reanalysis 1957-2001. ERA-40 Project Series 3 Workshop on Reanalysis. ECMWF, Shinfield Park, Reading RG2 9AX.